A M E N D M E N T S

In the Specification:

Please replace the Brief Description of the Drawings section with the following replacement section:

Brief Description of the Drawings

FIG. 1 schematically shows an optimized construction of a synthetic gene coding for hG-CSF according to a preferred embodiment of the present invention.

FIG. 2 shows SEQ ID NO:3, the DNA sequence of the native gene sequence coding for hG-CSF (FIG. 2A) (GenBank: NM _ 000759) and SEQ ID NO:1, the DNA sequence of the optimized (Fopt5) gene coding for hG-CSF (FIG. 2B). The bases which differ from native gene are bolded.

FIG. 3 shows an SDS-PAGE analysis of samples of proteins obtained from the expression of native hG-CSF DNA sequence (lanes 1 to 4) and of optimized (Fopt5) gene coding for hG-CSF (lanes 6 and 7) in induced and noninduced cultures of *E. coli* , as evaluated by dye staining (FIG. 3A) and by Western blot using antibody specific for hG-CSF protein (FIG. 3B).
FIG. 4 shows an SDS-PAGE analysis of samples of proteins obtained from the expression of optimized (Fopt5) gene coding for hG-CSF in induced culture of *E. coli* , as evaluated by dye staining.

FIG. 5 shows an SDS-PAGE analysis of samples of proteins obtained from the expression of optimized (Fopt5) gene coding for hG-CSF in induced culture of *E. coli* according to an alternative embodiment, as evaluated by dye staining.

Please replace the first full paragraph on page 6 of the specification with the following replacement paragraph:

It has been found that the problem with the low expression level of the gene coding for hG-CSF in *E. coli* can be solved by the optimization of the gene sequence coding for hG-CSF. The native gene coding for hG-CSF, defined by SEQ ID NO:3, is changed, leading to the construction of a particular synthetic gene coding for hG-CSF. The particular synthetic gene is defined by the DNA sequence of SEQ ID NO:1 or by a nucleotide sequence comprising suitable modifications of SEQ ID NO:1 or of the native hG-CSF gene sequence, SEQ ID NO:3.

Please replace the seventh full paragraph on page 7 of the specification with the following replacement paragraph:

The term 'segment I', as used herein, refers to the 5′ end of the gene encoding hG-CSF, defined by SEQ ID NOS: 1, 2, or 3, between the nucleotide positions 3 and 194 (notably the restriction sites NdeI (3) and SacI (194)), i.e. 191 bp long sequence. Segment I may be de novo synthesized.

Please replace the eighth full paragraph on page 7 of the specification with the following replacement paragraph:

The term 'segment II', as used herein, refers to the part of the gene for hG-CSF, defined by SEQ ID NOS: 1, 2, or 3, between the nucleotide positions 194 and 309 (notably the restriction sites SacI (194) and ApaI (309)), i.e. 115 bp long central part of the gene. Segment II may be de novo synthesized.

Please replace the paragraph bridging pages 7 and 8 of the specification with the following replacement paragraph:

The term 'segment III', as used herein, refers to the part of the gene for hG-CSF, defined by SEQ ID NOS: 1, 2, or 3, between the nucleotide positions 309 and 467 (notably the restriction sites ApaI (309) and NheI (467)), i.e. 158 bp long part of the gene where the native DNA sequence for hG-CSF is preserved with the exception of codons for Arg148 and Gly150.

Please replace the first full paragraph on page 8 of the specification with the following replacement paragraph:

The term 'segment IV', as used herein, refers to the 3′ terminal end of the gene encoding hG-CSF, defined by SEQ ID NOS: 1, 2, or 3, between the nucleotide positions 467 and 536 (notably the restriction sites NheI (467) and BamHI (536)), i.e. 69 bp long terminal part of the gene. Segment IV may be de novo synthesized.

Please replace the second full paragraph on page 8 of the specification with the following replacement paragraph:

The synthetic gene encoding hG-CSF, one embodiment of which is defined by SEQ ID NO:1, of the present invention is constructed by the combination of the following methods:

- replacement of *E. coli* rare codons with *E. coli* preference codons: in the segment II (between restriction sites SacI (194) and ApaI (309)) and in the segment IV (between restriction sites NheI (467) and BamHI (536)) of SEQ ID NO: 3.
- replacement of GC rich regions with AT rich regions, whereby the rarest *E. coli* codons are replaced, but mostly not with the *E. coli* preference codons: in the segment I (between restriction sites NdeI (3) and SacI (194)) of SEQ ID NO:3.
- completely unchanged native sequence of 46 codons (between CCC for Pro102 and CGC for Arg147) in the segment III of SEQ ID NO:3.
- replacement of two *E. coli* rare codons (CGG→CGT (Arg148) and GGA→GGT (Gly150)) at the terminal end of the segment III of SEQ ID NO:3.

Please replace the fourth full paragraph on page 8 of the specification with the following replacement paragraph:

The synthetic gene of the present invention encoding hG-CSF, defined by SEQ ID NO:1, enables expression of the constructed synthetic gene encoding hG-CSF with the expression level in *E. coli* equal to or higher than 52%. Furthermore, the expression level of about 55% or even about 60% can also be obtained. High expression level of the synthetic gene coding

for hG-CSF, defined by SEQ ID NO:1, of the present invention enables high yields of hG-CSF production, faster and simpler purification and isolation of heterologous hG-CSF, easier in-process control, and the whole production process is more economical. Therefore, the efficient production of hG-CSF in industrial scale is enabled. The produced hG-CSF is suitable for clinical use in medicine.

Please replace the paragraph bridging pages 8 and 9 of the specification with the following replacement paragraph:

The construction of the synthetic gene of the present invention, defined by SEQ ID NO:1, begins with the initial preparation of the hG-CSF native gene and of the plasmids. Gene coding for native hG-CSF can be of human origin, such as the gene defined by SEQ ID NO:3, but the same principle can be used for every gene which is homologous in the regions which comprise single restriction sites which are used for subcloning of de novo synthesized gene segments. The plasmid for mutagenesis was chosen according to its ability to be capable of enabling the successive introduction of point mutations. Selection or enrichment of the plasmids containing desired mutation was obtained by using an additional selection primer that changed unique restriction site EcoRI into EcoRV or vice-versa (Transformer™ Site-Directed Mutagenesis Kit (Clontech)). The gene and the plasmid are constructed in such a way that the introduction of point mutation by cassette mutagenesis is possible.

Please replace the first full paragraph on page 9 of the specification with the following replacement paragraph:

After the initial preparation of native gene coding for hG-CSF, defined by SEQ ID NO:3, and of plasmids the optimization of the native gene coding for hG-CSF is performed. This means that the synthetic gene coding for hG-CSF, defined by SEQ ID NO:1, is constructed. The optimization begins with the division of the native gene coding for hG-CSF (SEQ ID NO:3) into four (I, II, III ~~in~~ and IV) segments, which are or will be separated with single restriction sites after the oligonucleotide mutagenesis and in the individual segments the changes are introduced. In some individual segments the changes in the gene sequence are introduced whereas in certain segments the gene is not changed (FIG. 1). The obtained optimized synthetic gene coding for hG-CSF (SEQ ID NO:1) therefore consists of partially preserved

native sequence (segment III) of SEQ ID NO:3 and of 5′ and 3′ coding regions which are synthesized de novo (segments I, II and IV).

Please replace the first full paragraph on page 13 of the specification with the following replacement paragraph:

The gene coding for hG-CSF, defined by SEQ ID NO:3, was amplified from BBG13 (R&D) with the PCR method, which was also used to introduce by using the start oligonucleotides the restriction sites NdeI and BamHI at the start and terminal end of the gene. The gene was then incorporated in the plasmid pCytexΔH,H (see the description below) between the restriction sites NdeI and BamHI. All other optimization steps for the expression of the gene in *E. coli* were also performed in this plasmid.

Please replace the second full paragraph on page 13 of the specification with the following replacement paragraph:

During the initial gene preparation the EcoRV restriction site was annihilated (oligo M20z108, defined by SEQ ID NO:4) by point mutation. This was performed with the aim to ensure the possibility of introduction of (individual) mutations by using the oligonucleotide-directed mutagenesis in the plasmid pCytexΔH,H with the kit Transformer™ Site-Directed Mutagenesis Kit (Clontech). The selection of mutants in the plasmid pCytexΔH,H-G-CSF via the restriction sites EcoRI/EcoRV was therefore possible.

Please replace the fourth full paragraph on page 13 of the specification with the following replacement paragraph:

The oligonucleotide for the annihilation of EcoRV site from the gene coding for hG-CSF:

M20z108 5′-CCT GGA AGG AAT ATC CCC CG-3′ (SEQ ID NO:4)

Please replace the paragraph bridging pages 13 and 14 of the specification with the following replacement paragraph:

In the first optimization step the synthetic gene between the restriction sites NdeI and SacI was constructed by ligation of five cassettes (A, B, C, D, E) which were composed of complementary oligonucleotides. This synthetic part of the gene represents the segment I.

With the segment I the part of the native gene for hG-CSF (SEQ ID NO:3) between the restriction sites NdeI and SacI was replaced. This was performed by the excision of the first part of the gene between the restriction sites NdeI and SacI and its replacement with the synthetically prepared cassette. The process was performed in two steps. In the first step, the cassette A, defined by SEQ ID NOS: 5 and 6, was ligated to the NdeI site and the cassette E, defined by SEQ ID NOS: 13 and 14, was ligated to the SacI site. After 16 hours at 16° C. the ligation mixture was precipitated with ethanol to remove the excess of (not bound) oligonucleotides. In the second steps the central part of the whole cassette (cassettes B, defined by SEQ ID NOS: 7 and 8, C, defined by SEQ ID NOS: 9 and 10, and D, defined by SEQ ID NOS: 11 and 12,) from the three previously ligated complementary oligonucleotides was added and the ligation was performed for 16 hours at 16° C.

Please replace the second full paragraph on page 14 of the specification with the following replacement paragraph:

In the third optimization step the segment IV was constructed in a similar way as the segment I with the exception of intermediate ethanol precipitation. The segment IV represents the last part of the gene between the restrictions sites NheI and BamHI and is composed of two pairs of complementary oligonucleotides (cassettes F, defined by SEQ ID NOS: 16 and 17, and G, defined by SEQ ID NOS: 18 and 19).

Please replace the fourth full paragraph on page 14 of the specification with the following replacement paragraph:

ApaI restriction site was then used in the fifth optimization step with the aim to replace the native gene between SacI and ApaI with the synthetic DNA (segment II). This synthetic DNA is composed of three pairs of complementary oligonucleotides (cassette H, defined by SEQ ID NOS: 21 and 22, I, defined by SEQ ID NOS: 23 and 24, and J, defined by SEQ ID NOS: 25 and 26). This was performed similarly as in the first step with the later addition of the cassette I.

Please replace the last full paragraph on page 14 through the last full paragraph on page 16 of the specification with the following replacement paragraphs:

1st Optimization Step:

Complementary pairs of oligonucleotides (NdeI-SacI; Segment I in FIG. 1):

Cassette A: Composed of complementary oligonucleotides zg1os1 in sp1os2:

zg1os1 5′ TAT GAC ACC ACT GGG TCC AGC TTC TTC TCT GCC GCA AAG 3′ (SEQ ID NO:5)

sp1os2 5′ GCA GAG AAG AAG CTG GAC CCA GTG GTG TCA 3′ (SEQ ID NO:6)

Cassette B: Composed of complementary oligonucleotides zg2os3 in sp2os4:

zg2os3 5′ CTT TCT GTT GAA ATG TTT AGA ACA AGTTCG TAA AAT TCA AG 3′ (SEQ ID NO:7)

sp2os4 5′ GAA CTT GTT CTA AAC ATT TCA ACA GAA AGC TTT GCG 3′ (SEQ ID NO:8)

Cassette C: Composed of complementary oligonucleotides zg3os5 in sp3os6:

zg3os5 5′ GTG ATG GTG CAG CTT TAC AAG AAA AAC TGT GTG 3′ (SEQ ID NO:9)

sp3os6 5′ GTT TTT CTT GTA AAG CTG CAC CAT CAC CTT GAA TTT TAC 3′ (SEQ ID NO:10)

Cassette D: Composed of complementary oligonucleotides zg4os7 in sp4os8:

zg4os7 5′ CAA CTT ATA AAC TGT GTC ATC CAG AAG AAC TGG TTC TGT TAG 3′ (SEQ ID NO:11)

sp4os8 5′ CAG TTC TTC TGG ATG ACA CAG TTT ATA AGT TGC ACA CA 3′ (SEQ ID NO:12)

Cassette E: Composed of complementary oligonucleotides zg5os9 in sp5os10:

zg5os9 5′ GTC ATT CTC TGG GTA TTC CGT GGG CTC CTC TGA GCT 3′ (SEQ ID NO:13)

sp5os10 5′ CAG AGG AGC CCA CGG AAT ACC CAG AGA ATG ACC TAA CAG AAC 3′ (SEQ ID NO:14)

2nd Optimization Step: Oligonucleotides for the replacement of the most critical codons by using the oligonucleotide-directed mutagenesis
replacement CGG→CGT (Arg 148) and GGA→GGT (Gly 150)
m38os16 5′ CTC TGC TTT CCA GCG CCG TGC AGG TGG GGT CCT GGT TG 3′ (SEQ ID NO:15)

3rd Optimization Step: Complementary pairs of nucleotides (NheI-BamHI; Segment IV on FIG. 1):
Cassette F: Composed of complementary nucleotides zg6os11 in sp6os12:
zg6os11 5′ CTA GCC ATC TGC AAT CTT TTC TGG AAG TTA G 3′ (SEQ ID NO:16)
sp6os12 5′ ACG ATA GCT AAC TTC CAG AAA AGA TTG CAG ATG G 3′ (SEQ ID NO:17)

Cassette G: Composed of complementary oligonucleotides zg7os13 in sp7os14:
zg7os13 5′ CTA TCG TGT TCT GCG TCA TCT GGC TCA GCC GTG ATA AG 3′ (SEQ ID NO:18)
sp7os14 5′ GAT CCT TAT CAC CGC TGA GCC AGA TGA CGC AGA AC 3′ (SEQ ID NO:19)

4th Optimization Step: Oligonucleotides for the introduction of ApaI (309) (GGT→GGG (Gly101)), and the replacement of the rare codon ATA→ATT (Ile96) by using the oligonucleotide-directed mutagenesis
insertion of ApaI (309) (GGT→GGG (Gly101)), and replacement ATA→ATT (Ile 96):
Apalos15 5′ GCC CTG GAG GGG ATT TCC CCC GAG TTG GGG CCC ACC TTG GAC AC 3′ (SEQ ID NO:20)

5th Optimization Step: Complementary pairs of oligonucleotides (SacI-ApaI; Segment II in FIG. 1):
Cassette H: Composed of complementary oligonucleotides zg8os18 in sp8os19:
zg8os18 5′ CCT GTC CGA GCC AGG CGC TGC AGC TGG CAG GCT CCC TGA G 3′ (SEQ ID NO:21)

sp8os19 5′ CCT GCC AGC TGC AGC GCC TGG CTC GGA CAG GAG CT 3′ (SEQ ID NO:22)

Cassette I: Composed of complementary oligonucleotides zg9os20 in sp9os21:

zg9os20 5′ CCA ACT GCA TAG CGG TCT GTT TCT GTA TCA GGG TCT GCT G 3’ (SEQ ID NO:23)

sp9os21 5′ CTG ATA CAG AAA CAG ACC GCT ATG CAG TTG GCT CAG GCA G 3’ (SEQ ID NO:24)

Cassette J: Composed of complementary oligonucleotides zg10os22 in sp10os23:

zg10os22 5′ CAG GCG CTG GAA GGC ATT TCC CCG GAA CTG GGG CC 3′ (SEQ ID NO:25)

sp10os23 5′ CCA GTT CCG GGG AAA TGC CTT CCA GCG CCT GCA GCA GAC C 3′ (SEQ ID NO:26)

Please replace the paragraph bridging pages 16 and 17 of the specification with the following replacement paragraph:

The optimized gene Fopt5, defined by SEQ ID NO:1, was excised from the plasmid pCyΔH,H with the restriction enzymes NdeI and BamHI and the gene was then subcloned in the final expression plasmid pET3a (Novagen, Madison USA), which contains an ampicilline resistance gene, which was then transformed into the production strain *E. coli* BL21 DE3).

Please replace the Sequence Listing Section with the following replacement section:

SEQUENCE LISTING

<110> Lek Pharmaceuticals d. d.

<120> Synthetic gene coding for human granulocyte-colony stimulating factor for the expression in E. coli

<130> 32992

<140> US/10/522,827A
<141> 2005-01-31

<160> 26

<170> PatentIn version 3.5

<210> 1
<211> 525
<212> DNA
<213> artificial

<220>
<223> synthetic

<400> 1
atgacaccac tgggtccagc ttcttctctg ccgcaaagct ttctgttgaa atgtttagaa 60

caagttcgta aaattcaagg tgatggtgca gctttacaag aaaaactgtg tgcaacttat 120

aaactgtgtc atccagaaga actggttctg ttaggtcatt ctctgggtat tccgtgggct 180

cctctgagct cctgtccgag ccaggcgctg cagctggcag gctgcctgag ccaactgcat 240

agcggtctgt ttctgtatca gggtctgctg caggcgctgg aaggcatttc cccggaactg 300

gggcccacct tggacacact gcagctggac gtcgccgact ttgccaccac catctggcag 360

cagatggaag aactgggaat ggcccctgcc ctgcagccca cccagggtgc catgccggcc 420

ttcgcctctg ctttccagcg ccgtgcaggt ggggtcctgg ttgctagcca tctgcaatct 480

tttctggaag ttagctatcg tgttctgcgt catctggctc agccg 525

<210> 2
<211> 528
<212> DNA
<213> artificial

<220>
<223> synthetic

<400> 2
atgacaccac tgggtccagc ttcttctctg ccgcaaagct ttctgttgaa atgtttagaa 60

caagttcgta aaattcaagg tgatggtgca gctttacaag aaaaactgtg tgcaacttat 120

aaactgtgtc atccagaaga actggttctg ttaggtcatt ctctgggtat tccgtgggct 180

cctctgagct cctgtccgag ccaggcgctg cagctggcag gctgcctgag ccaactgcat 240

agcggtctgt ttctgtatca gggtctgctg caggcgctgg aaggcatttc cccggaactg 300

gggcccacct tggacacact gcagctggac gtcgccgact ttgccaccac catctggcag 360

cagatggaag aactgggaat ggcccctgcc ctgcagccca cccagggtgc catgccggcc 420

ttcgcctctg ctttccagcg ccgtgcaggt ggggtcctgg ttgctagcca tctgcaatct 480

tttctggaag ttagctatcg tgttctgcgt catctggctc agccgtga 528

<210> 3
<211> 525
<212> DNA
<213> Homo sapiens

<400> 3
atgacccccc tgggccctgc cagctccctg ccccagagct tcctgctcaa gtgcttagag 60

caagtgagga agatccaggg cgatggcgca gcgctccagg agaagctgtg tgccacctac 120

aagctgtgcc accccgagga gctggtgctg ctcggacact ctctgggcat cccctgggct 180

cccctgagct cctgccccag ccaggccctg cagctggcag gctgcttgag ccaactccat 240

agcggccttt tcctctacca ggggctcctg caggccctgg aagggatatc ccccgagttg 300

ggtcccacct tggacacact gcagctggac gtcgccgact ttgccaccac catctggcag 360

cagatggaag aactgggaat ggcccctgcc ctgcagccca cccagggtgc catgccggcc 420

ttcgcctctg ctttccagcg ccgggcagga ggggtcctgg ttgctagcca tctgcagagc 480

ttcctggagg tgtcgtaccg cgttctacgc caccttgcgc agccc 525

<210> 4
<211> 20
<212> DNA
<213> artificial

<220>
<223> synthetic

<400> 4
cctggaagga atatcccccg 20

<210> 5
<211> 39
<212> DNA
<213> artificial

<220>
<223> synthetic

<400> 5
tatgacacca ctgggtccag ctccttctct gccgcaaag 39

<210> 6
<211> 30

<212> DNA

<213> Artificial

<220>

<223> synthetic

<400> 6

gcagagaaga agctggaccc agtggtgtca 30

<210> 7

<211> 41

<212> DNA

<213> artificial

<220>

<223> synthetic

<400> 7

ctttctgttg aaatgtttag aacaagttcg taaaattcaa g 41

<210> 8

<211> 36

<212> DNA

<213> artificial

<220>

<223> synthetic

<400> 8

gaacttgttc taaacatttc aacagaaagc tttgcg 36

<210> 9
<211> 33
<212> DNA
<213> artificial

<220>
<223> synthetic

<400> 9
gtgatggtgc agctttacaa gaaaaactct gtg 33

<210> 10
<211> 39
<212> DNA
<213> artificial

<220>
<223> synthetic

<400> 10
gtttttcttg taaagctgca ccatcacctt gaattttac 39

<210> 11
<211> 42
<212> DNA
<213> artificial

<220>

<223> synthetic

<400> 11

caacttataa actgtgtcat ccagaagaac tggttctgtt ag 42

<210> 12

<211> 38

<212> DNA

<213> artificial

<220>

<223> synthetic

<400> 12

cagttcttct ggatgacaca gtttataagt tgcacaca 38

<210> 13

<211> 36

<212> DNA

<213> artificial

<220>

<223> synthetic

<400> 13

gtcattctct gggtattccg tgggctcctc tgagct 36

<210> 14

<211> 42

<212> DNA

<213> artificial

<220>

<223> synthetic

<400> 14

cagaggagcc cacggaatac ccagagaatg acctaacaga ac 42

<210> 15

<211> 38

<212> DNA

<213> artificial

<220>

<223> synthetic

<400> 15

ctctgctttc cagcgccgtg caggtggggt cctggttg 38

<210> 16

<211> 31

<212> DNA

<213> artificial

<220>

<223> synthetic

<400> 16
ctagccatct gcaatccttt ctggaagtta g 31

<210> 17
<211> 34
<212> DNA
<213> artificial

<220>
<223> synthetic

<400> 17
acgatagcta acttccagaa aagattgcag atgg 34

<210> 18
<211> 38
<212> DNA
<213> artificial

<220>
<223> synthetic

<400> 18
ctatcgtgtt ctgcgtcatc tggctcagcc gtgataag 38

<210> 19

<211> 35

<212> DNA

<213> artificial

<220>

<223> synthetic

<400> 19

gatccttatc acggctgagc cagatgacgc agaac 35

<210> 20

<211> 44

<212> DNA

<213> artificial

<220>

<223> synthetic

<400> 20

gccctggagg ggatttcccc cgagttgggg cccaccttgg acac 44

<210> 21

<211> 40

<212> DNA

<213> artificial

<220>

<223> synthetic

<400> 21

cctgtccgag ccaggcgctg cagctggcag gctgcctgag 40

<210> 22

<211> 35

<212> DNA

<213> artificial

<220>

<223> synthetic

<400> 22

cctgccagct gcagcgcctg gctcggacag gagct 35

<210> 23

<211> 40

<212> DNA

<213> artificial

<220>

<223> synthetic

<400> 23

ccaactgcat agcggtctgt ttctgtatca gggtctgctg 40

<210> 24

<211> 40

<212> DNA

<213> artificial

<220>

<223> synthetic

<400> 24

ctgatacaga aacagaccgc tatgcagttg gctcaggcag 40

<210> 25

<211> 35

<212> DNA

<213> artificial

<220>

<223> synthetic

<400> 25

caggcgctgg aaggcatttc cccggaactg gggcc 35

<210> 26

<211> 40

<212> DNA

<213> artificial

<220>

<223> synthetic

<400> 26

ccagttccgg ggaaatgcct tccagcgcct gcagcagacc 40